

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/426,509DATE: 08/28/95
TIME: 13:46:12

INPUT SET: S5863.raw

This Raw Listing contains the General
Information Section and up to the first 5 pages.

SEQUENCE LISTING

ENTERED

(1) General Information:

(i) APPLICANT: Ullrich, Axel
Gishizsky, Mikhail
Sures, Irman G.

(ii) TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
Kinases

(iii) NUMBER OF SEQUENCES: 21

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Pennie & Edmonds
(B) STREET: 1155 Avenue of the Americas
(C) CITY: New York
(D) STATE: New York
(E) COUNTRY: U.S.A.
(F) ZIP: 10036

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/426,509
(B) FILING DATE: 21-APR-1995
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 08/232,545
(B) FILING DATE: 22-APR-1994
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Coruzzi, Laura A.
(B) REGISTRATION NUMBER: 30,742
(C) REFERENCE/DOCKET NUMBER: 7683-074

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (212)790-9090
(B) TELEFAX: (212)869-9741
(C) TELEX: 66141 PENNIE

RAW SEQUENCE LISTING
PATENT APPLICATION US/08/426,509DATE: 08/28/95
TIME: 13:46:15

INPUT SET: S5863.raw

47
48
49 (2) INFORMATION FOR SEQ ID NO:1:
50
51 (i) SEQUENCE CHARACTERISTICS:
52 (A) LENGTH: 2000 base pairs
53 (B) TYPE: nucleic acid
54 (C) STRANDEDNESS: unknown
55 (D) TOPOLOGY: unknown
56
57 (ii) MOLECULE TYPE: DNA
58
59
60
61
62 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
63
64 CTCGCTCCAA GTTGTGCAGC CGGGACCGCC TCGGGGTGTG CAGCCGGCTC GCGGAGGCCC 60
65
66 TCCTGGGGGC GGGCGCGGGG CGGCTCGGGG GCGCCCCCTG AGCAGAAAAC AGGAAGAACC 120
67
68 AGGCTCGGTC CAGTGGCACC CAGCTCCCTA CCTCCTGTGC CAGCCGCCTG GCCTGTGGCA 180
69
70 GGCCATTCCC AGCGTCCCCG ACTGTGACCA CTTGCTCAGT GTGCCCTCTCA CCTGCCCTCAG 240
71
72 TTTCCCTCTG GGGGGCGATG GCGGGGCGAG GCTCTCTGGT TTCTTGCGCG GCATTTTCACG 300
73
74 GCTGTGATTC TGCTGAGGAA CTTCCCCGGG TGAGCCCCCG CTTCCTCCGA GCCTGGCACC 360
75
76 CCCCTCCCGT CTCAGCCAGG ATGCCAACGA GGCCTGGGC CCCGGGCACC CAGTGTATCA 420
77
78 CCAAATGCGA GCACACCCGC CCAAGCCAG GGGAGCTGGC CTTCGCAAG GGCACGTGG 480
79
80 TCACCATCCT GGAGGCCTGC GAGAACAAGA GCTGGTACCG CGTCAAGCAC CACACCAGTG 540
81
82 GACAGGAGGG GCTGCTGGCA GCTGGGGCGC TGCGGGAGCG GGAGGCCCTC TCCGCAGACC 600
83
84 CCAAGCTCAG CCTCATGCCG TGGTTCCACG GGAAGATCTC GGGCCAGGAG GCTGTCCAGC 660
85
86 AGCTGCAGCC TCCCGAGGAT GGGCTGTTCC TGGTGCGGGA GTCCGCGCGC CACCCCGGCG 720
87
88 ACTACGTCCT GTGCGTGAGC TTTGGCCGCG ACGTCATCCA CTACCGCGTG CTGCACCGCG 780
89
90 ACGGCCACCT CACAATCGAT GAGGCCGTGT TCTTCTGCAA CCTCATGGAC ATGGTGGAGC 840
91
92 ATTACAGCAA GGACAAGGGC GCTATCTGCA CCAAGCTGGT GAGACCAAAG CGGAAACACG 900
93
94 GGACCAAGTC GGCCGAGGAG GAGCTGGCCA GGGCGGGCTG GTTACTGAAC CTGCAGCATT 960
95
96 TGACATTGGG AGCACAGATC GGAGAGGGAG AGTTTGGAGC TGTCTGCAG GGTGAGTACC 1020
97
98 TGGGGCAAAA GGTGGCCGTG AAGAATATCA AGTGTGATGT GACAGCCCAG GCCTTCCTGG 1080
99

RAW SEQUENCE LISTING PATENT APPLICATION US/08/426,509

DATE: 08/28/95
TIME: 13:46:18

INPUT SET: S5863.raw

```

100  ACGAGACGGC CGTCATGACG AAGATGCAAC ACGAGAACCCT GGTGCGTCTC CTGGGCGTGA      1140
101
102  TCCTGCACCA GGGGCTGTAC ATTGTCATGG AGCACGTGAG CAAGGGCAAC CTGGTGAAC      1200
103
104  TTCTGCGGAC CCGGGGTCGA GCCCTCGTGA ACACCGCTCA GCTCCTGCAG TTTTCTCTGC      1260
105
106  ACGTGGCCGA GGGCATGGAG TACCTGGAGA GCAAGAAGCT TGTGCACCGC GACCTGGCCG      1320
107
108  CCCGCAACAT CCTGGTCTCA GAGGACCTGG TGGCCAAGGT CAGCGACTTT GGCCTGGCCA      1380
109
110  AAGCCGAGCG GAAGGGGCTA GACTCAAGCC GGCTGCCCCG CAAGTGGACG GCGCCCAGAG      1440
111
112  CTCTCAAACA CGGGAAGTTC ACCAGCAAGT CGGATGTCTG GAGTTTTGGG GTGCTGCTCT      1500
113
114  GGGAGGTCTT CTCATATGGA CGGGCTCCGT ACCCTAAAAT GTCACTGAAA GAGGTGTCGG      1560
115
116  AGGCCGTGGA GAAGGGGTAC CGCATGGAAC CCCCCGAGGG CTGTCCAGGC CCCGTGCACG      1620
117
118  TCCTCATGAG CAGCTGCTGG GAGGCAGAGC CCGCCCGCCG GCCACCCTTC CGCAAAC      1680
119
120  CCGAGAAGCT GGCCCGGGAG CTACGCAGTG CAGGTGCCCC AGCCTCCGTC TCAGGGCAGG      1740
121
122  ACGCCGACGG CTCCACCTCG CCCCAGAGCC AGGAGCCCTG ACCCCACCCG GTGGGGCCCT      1800
123
124  TGGCCCCAGA GGACCGAGAG AGTGGAGAGT GCGGCGTGGG GGCAC      1860
125
126  AGGGTCCAGG CGGGCAAGTC ATCCTCCTGG TGCCACAGC AGGGGCTGGC CCACGTAGGG      1920
127
128  GGCTCTGGGC GGCCCGTGGA CACCCAGAC CTGCGAAGGA TGATCGCCCG ATAAAGACGG      1980
129
130  ATTCTAAGGA CTCTAAAAAA      2000
131

```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 507 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

Met Ala Gly Arg Gly Ser Leu Val Ser Trp Arg Ala Phe His Gly Cys
1           5           10           15
Asp Ser Ala Glu Glu Leu Pro Arg Val Ser Pro Arg Phe Leu Arg Ala
20           25           30
Trp His Pro Pro Pro Val Ser Ala Arg Met Pro Thr Arg Arg Trp Ala

```

RAW SEQUENCE LISTING PATENT APPLICATION US/08/426,509

DATE: 08/28/95
TIME: 13:46:20

INPUT SET: S5863.raw

	35	40	45
153			
154			
155	Pro Gly Thr Gln Cys Ile	Thr Lys Cys Glu His	Thr Arg Pro Lys Pro
156	50	55	60
157			
158	Gly Glu Leu Ala Phe Arg	Lys Gly Asp Val Val	Thr Ile Leu Glu Ala
159	65	70	75 80
160			
161	Cys Glu Asn Lys Ser Trp	Tyr Arg Val Lys His	His Thr Ser Gly Gln
162		85	90 95
163			
164	Glu Gly Leu Leu Ala Ala	Gly Ala Leu Arg Glu	Arg Glu Ala Leu Ser
165		100	105 110
166			
167	Ala Asp Pro Lys Leu Ser	Leu Met Pro Trp Phe	His Gly Lys Ile Ser
168		115	120 125
169			
170	Gly Gln Glu Ala Val Gln	Gln Leu Gln Pro Pro	Glu Asp Gly Leu Phe
171		130	135 140
172			
173	Leu Val Arg Glu Ser Ala	Arg His Pro Gly Asp	Tyr Val Leu Cys Val
174	145	150	155 160
175			
176	Ser Phe Gly Arg Asp Val	Ile His Tyr Arg Val	Leu His Arg Asp Gly
177		165	170 175
178			
179	His Leu Thr Ile Asp Glu	Ala Val Phe Phe Cys	Asn Leu Met Asp Met
180		180	185 190
181			
182	Val Glu His Tyr Ser Lys	Asp Lys Gly Ala Ile	Cys Thr Lys Leu Val
183		195	200 205
184			
185	Arg Pro Lys Arg Lys His	Gly Thr Lys Ser Ala	Glu Glu Glu Leu Ala
186		210	215 220
187			
188	Arg Ala Gly Trp Leu Leu	Asn Leu Gln His Leu	Thr Leu Gly Ala Gln
189	225	230	235 240
190			
191	Ile Gly Glu Gly Glu Phe	Gly Ala Val Leu Gln	Gly Glu Tyr Leu Gly
192		245	250 255
193			
194	Gln Lys Val Ala Val Lys	Asn Ile Lys Cys Asp	Val Thr Ala Gln Ala
195		260	265 270
196			
197	Phe Leu Asp Glu Thr Ala	Val Met Thr Lys Met	Gln His Glu Asn Leu
198		275	280 285
199			
200	Val Arg Leu Leu Gly Val	Ile Leu His Gln Gly	Leu Tyr Ile Val Met
201		290	295 300
202			
203	Glu His Val Ser Lys Gly	Asn Leu Val Asn Phe	Leu Arg Thr Arg Gly
204	305	310	315 320
205			

RAW SEQUENCE LISTING PATENT APPLICATION US/08/426,509

DATE: 08/28/95
TIME: 13:46:23

INPUT SET: S5863.raw

```

206      Arg Ala Leu Val Asn Thr Ala Gln Leu Leu Gln Phe Ser Leu His Val
207                      325                      330                      335
208
209      Ala Glu Gly Met Glu Tyr Leu Glu Ser Lys Lys Leu Val His Arg Asp
210                      340                      345                      350
211
212      Leu Ala Ala Arg Asn Ile Leu Val Ser Glu Asp Leu Val Ala Lys Val
213                      355                      360                      365
214
215      Ser Asp Phe Gly Leu Ala Lys Ala Glu Arg Lys Gly Leu Asp Ser Ser
216                      370                      375                      380
217
218      Arg Leu Pro Val Lys Trp Thr Ala Pro Glu Ala Leu Lys His Gly Lys
219                      385                      390                      395                      400
220
221      Phe Thr Ser Lys Ser Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu
222                      405                      410                      415
223
224      Val Phe Ser Tyr Gly Arg Ala Pro Tyr Pro Lys Met Ser Leu Lys Glu
225                      420                      425                      430
226
227      Val Ser Glu Ala Val Glu Lys Gly Tyr Arg Met Glu Pro Pro Glu Gly
228                      435                      440                      445
229
230      Cys Pro Gly Pro Val His Val Leu Met Ser Ser Cys Trp Glu Ala Glu
231                      450                      455                      460
232
233      Pro Ala Arg Arg Pro Pro Phe Arg Lys Leu Ala Glu Lys Leu Ala Arg
234                      465                      470                      475                      480
235
236      Glu Leu Arg Ser Ala Gly Ala Pro Ala Ser Val Ser Gly Gln Asp Ala
237                      485                      490                      495
238
239      Asp Gly Ser Thr Ser Pro Arg Ser Gln Glu Pro
240                      500                      505
241

```

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2500 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: unknown
- (D) TOPOLOGY: unknown

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCGCTTTTTTG CTTAGAGCTT GAGAGTCAAA GTTAAGGACC CACATGTATA CTTCGGCTCT

60

PAGE: 1

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION US/08/426,509

DATE: 08/28/95
TIME: 13:46:26

INPUT SET: S5863.raw

Line

Error

Original Text

NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES

The nucleotide and/or amino acid sequence disclosure contained in this application does not comply with the requirements for such a disclosure as set forth in 37 CFR 1.821 - 1.825 for the following reason(s):

- ☒ 1. This application clearly fails to comply with the requirements of 37 CFR 1.821 - 1.825. Applicant's attention is directed to these regulations, published at 1114 OG 29, May 15, 1990 and at 55 FR 18230, May 1, 1990.
- ☐ 2. This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing" as required by 37 CFR 1.821(c).
- ☐ 3. A copy of the "Sequence Listing" in computer readable form has not been submitted as required by 37 CFR 1.821(e).
- ☐ 4. A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 CFR 1.822 and/or 1.823, as indicated on the attached marked-up copy of the "Raw Sequence Listing."
- ☐ 5. The computer readable form that has been filed with this application has been found to be damaged and/or unreadable as indicated on the attached CRF Diskette Problem Report. A substitute computer readable form must be submitted as required by 37 CFR 1.825(d).
- ☐ 6. The paper copy of the "Sequence Listing" is not the same as the computer readable form of the "Sequence Listing" as required by 37 CFR 1.821(e).
- ☐ 7. Other: _____

Applicant must provide:

- ☒ An initial or substitute computer readable form (CRF) copy of the "Sequence Listing"
- ☒ An initial or substitute paper copy of the "Sequence Listing", as well as an amendment directing its entry into the specification
- ☒ A statement that the content of the paper and computer readable copies are the same and, where applicable, include no new matter, as required by 37 CFR 1.821(e) or 1.821(f) or 1.821(g) or 1.825(b) or 1.825(d)

For questions regarding compliance with these requirements, please contact:

For Rules Interpretation, call (703) 308-1123
For CRF submission help, call (703) 308-4212
For PatentIn software help, call (703) 308-6856

Please return a copy of this notice with your response.